Massachusetts Institute of Technology | Handout 38
6.042J/18.062J: Mathematics for Computer Science | November 2, 1995
Professors David Karger and Tom Leighton

Notes 15

1 Permutations of Multisets

A multiset is a collection of objects, some of which may not be distinguishable from each other.

Example 1 $\{a, b, a, c, b, a\} = \{3 \cdot a, 2 \cdot b, 1 \cdot c\}$.

Since some objects are not distinguishable, there will be permutations that are indistinguishable because the only difference is in the positions of indistinguishable objects (e.g., an object and its "twin" swap places).

It is easy to see that for a multiset of k objects, each with infinite repetitions, we have k^r r-permutations. The problem is more interesting if we only allow a finite number of repetitions. This is the content of the following theorem:

Theorem 1.1 *Let S be a multiset with finite repetition counts $n_1, n_2, \ldots n_k$, where $n_1 + n_2 + \ldots + n_k = n = |S|$. Then the number of permutations of S is*

$$\frac{n!}{n_1! n_2! \cdots n_k!}.$$

Example 2 The number of distinguishable permutations of the letters of PEPPER is $\frac{6!}{3!2!1!} = 60$.

Proof. We prove our theorem by the above example. Subscript each letter to make them distinct. Then we have the letters $P_1, P_2, P_3, E_1, E_2, R_1$. This set has 6! permutations. Though, by mapping all the P_i to P we get a 3!:1 mapping, by mapping the E_i to E, we get a 2!:1 mapping, therefore, by the rule of division we get the statement of the theorem. ∎

We give another proof, this time *not* by example:

Proof. Let $S = \{n_1 \cdot a_1, n_2 \cdot a_2, \ldots, n_k \cdot a_k\}$. Consider n slots in which we put the elements (exactly one element to each slot) for each permutation. Then there are $\binom{n}{n_1}$ ways to choose slots for the a_1's, there are $\binom{n-n_1}{n_2}$ ways to put in the a_2's to the remaining slots, and so on. This shows that the total number of permutations is:

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n-n_1-\cdots-n_{k-1}}{n_k}=\frac{n!}{n_1!n_2!\cdots n_k!}$$

■

The quantity above is the *multinomial coefficient*

$$\binom{n_1+n_2+\ldots+n_k}{n_1,n_2,\ldots,n_k}$$

2 Combinations of Multisets

Let $S=\{2\cdot a, 1\cdot b, 3\cdot c\}$ The *2-combinations* of S are all the submultisets of S having two elements. That is, they are $\{2\cdot a\}, \{a,b\}, \{a,c\}, \{b,c\}, \{2\cdot c\}$.

Theorem 2.1 *The number of r-combinations of a multiset of k distinct objects, each with infinite repetition count is* $\binom{k+r-1}{r}$.

Remark. There are two other equivalent descriptions:

- The number of ways to put r indistinguishable balls in k bins
- The number of ways to write r as a sum of at most k positive numbers

■

Proof. Let $S=\{\infty\cdot a_1,\infty\cdot a_2,\infty\cdot a_k\}$. Then each combination can be written in the form

$$\{r_1\cdot a_1, r_2\cdot a_2,\ldots,r_k\cdot a_k\},$$

where $r_1+\cdots r_k=r$. Observe that there is a natural bijection from the set of the r-combinations of S written in canonical form and the set of arrangements of r $*$'s and k-1 bars in a line. Thus it is enough to compute the number of such arrangements, but this is clearly $\binom{k+r-1}{r}=\binom{k+r-1}{k-1}$ as all we need to do to characterize an arrangement is to tell where the $*$'s (or the bars) are. ■

Example 3 The number of nondecreasing 7-subsequences of 1,2,3,4,5,6,7,8,9,10 is $\binom{10+7-1}{7}=\binom{16}{7}$ as we can take any 7-combination of these ten numbers and arrange it in nondecreasing order to get such a sequence. How many increasing sequences are there? $\binom{10}{7}$.

Corollary 2.2 *The number of r-combinations of a multiset of $n \leq r$ objects, each with infinite repetition count, where each object appears at least once is $\binom{r-1}{n-1}$.*

Remark. Equivalent experiments: the number of ways to put r balls in n bins so that none is empty; the number of ways to write n as a sum of exactly r non-zero terms. ■

Proof. We know we have at least one of each object, so we really need to choose only r-n objects from the n. By the previous theorem we know that the number of ways to do that is

$$\binom{n-1+(r-n)}{r-n} = \binom{r-1}{r-n} = \binom{r-1}{n-1}$$

since $\binom{a}{b} = \binom{a}{a-b}$. ■

Example 4 In how many ways can the $2n+1$ seats in a congress be divided among three parties so that any two parties can form a majority?

To compute this number, we first compute the number of all assignments of the seats to the parties, without any restrictions. This is clearly $\binom{2n+3}{2}$ as we look for the number of (2n+1)-combinations of three objects with infinite counts. Second, we compute the number of *wrong* distributions, ie. when there is a party with at least n+1 seats. We have three choices for this unique party, then we can give it n+1 seats and distribute the remaining n seats in $\binom{n+2}{2}$ ways. Thus the number of wrong distributions is $3\binom{n+2}{2}$, therefore the number of good distributions is

$$\binom{2n+1}{2} - 3\binom{n+2}{2} = \frac{n(n+1)}{2} = \binom{n+1}{2}.$$

Example 5 What can we say if there are only $2n$ seats and thus the majority is $n+1$ people? We have to be more careful with this problem. We can certainly compute the number of all distributions as before to get $\binom{2n+2}{2}$. What about the number of wrong permutations? If the parties B and C fail to have absolute majority, then A has at least n seats. Thus we can give to A n seats in advance and then distribute the $n-1$ remaining seats to among the three parties in $\binom{n+2}{2}$ ways. The same can be said if B or C has at least n seats. However, we *cannot* say that "therefore, the number of wrong distributions is $3\binom{n+2}{2}$" as we deduced before, because three distributions, namely (n,n,0), (n,0,n) and (0,n,n) were counted twice. (This couldn't happen in the previous example since there couldn't be more than one party which held alone at least the half of the seats). Therefore, the number of wrong distributions is $3\binom{n+2}{2} - 3$, thus the number of good distributions is

$$\binom{2n+2}{2} - 3\binom{n+2}{2} - 3 = \frac{(n-1)(n-2)}{2} = \binom{n-1}{2}.$$

3 Physics

Statistical Mechanics studies the behavior of large systems of particles (e.g., boiling a kettle of water). One of the most basic physical systems is a *gas:* a collection of particles (electrons, atoms, photons) floating around in a large container.

One way to describe this system is to divide the container into lots of tiny boxes. We can then describe the system by saying, for each particle, which tiny box it is located in. We call this the "state" of the system. The system is equally likely to be in any one of its states (assuming they all have the same energy, but discussing this would take us too far afield). Many facts about the aggregate behavior of the system arise from averaging over all the possible states in the system.

A single particle is equally likely to be in any one of the tiny boxes. The obvious thing to do is to describe states by where particle 1 is, where particle 2 is, and so on. This means that if there are r particles and n positions, there are n^r possible states. This gives rise to the *Maxwell-Boltzman statistics* for the system. Nature follows these rules when all the particles are distinct (e.g, different atoms).

You would think it doesn't make a difference if the particles are all the same—after all, we could label them. Surprisingly, it matters! In a cloud of identical photons, nature actually obeys our balls in bins model. Thus there are $\binom{n+r-1}{n-1}$ possible states of the system, each equally likely. Gases made up of photons are said to obey *Bose-Einstein* statistics; photons are a kind of *Boson* and so is Helium 2 (the superconducting stuff). This has impacts on the physical behavior of the system. For example, suppose $r/n = \lambda$ is the average number of particles per cell. Under Maxwell-Boltzman Statistics, the fraction of empty cells is about

$$\lambda e^{-\lambda}$$

while under Bose–Einstein Statistics, the fraction of empty cells is about

$$\frac{1}{1+\lambda}$$

When λ is large, the first quantity is a lot less than the second. In other words, Maxwell-Boltzman predicts a more spread-out gas than you would actually find.

A different story happens with electrons, protons, and other "solid" particles called *Fermions.* At most one fermion will be present in a given cell (thus we'd better have $n > r$). If the particles were distinguishable, there would be $P(n, r) = n(n-1)\cdots(n-r+1)$ different states. But when the particles are all the same, nature acts as if there are $\binom{n}{r}$ different states. This means that gases made up of just one atom behave differently than gasses made up of more than one atom. In particular, the *entropy* of a gas, a physically measurable quantity, depends on the number of states. Physical experiments have shown that the entropy agrees with the indistinguishable particles model, and not the other.